

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-048019
(43)Date of publication of application : 18.02.2000

(51)Int.Cl. G06F 17/22
G06F 3/00
G06F 3/02

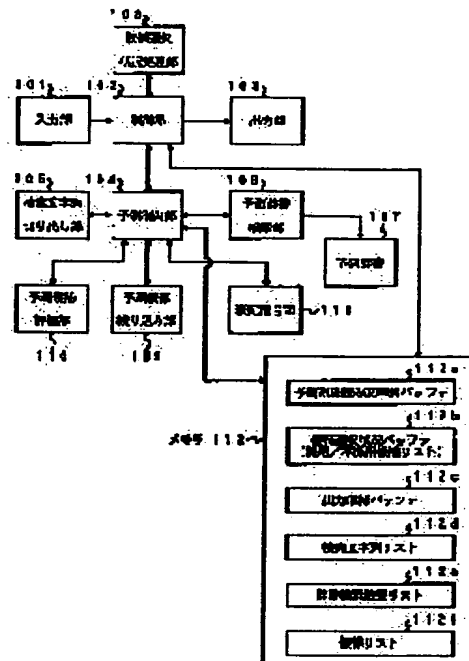
(21)Application number : 10-211118 (71)Applicant : TOSHIBA CORP
(22)Date of filing : 27.07.1998 (72)Inventor : ICHIMURA YUMI

(54) INPUT PREDICTION DEVICE, ITS METHOD AND RECORDING MEDIUM RECORDING INPUT PREDICTION PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To exclude unnecessary candidates and to provide only necessary candidates to a user by automatically executing input prediction on an optional position during the input of a character string and considering consistency with already displayed KANJI (Chinese character)/KANA (Japanese syllabary) notation and a user's adoption/non-adoption instruction.

SOLUTION: A character string inputted from an input part 101 is transferred to a retrieved character string segmentation part 105 to generate a partial character string group including a final character. The partial character string group is transferred to a prediction dictionary retrieving part 106 to obtain a predicted candidate, which is provided to an output part 103. When a provided predicted candidate is not adopted, a prediction control part 104 inhibits the output of the predicted candidate during the period of input prediction to be executed based on a character string including a character string referred to at the time of generating the predicted candidate. In addition, the output of a predicted candidate inconsistent with the notation of an already displayed character string is inhibited. Consequently unnecessary candidates are excluded and only necessary candidates can be provided to the user.



LEGAL STATUS

[Date of request for examination]
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2000-48019
(P2000-48019A)

(43)公開日 平成12年2月18日(2000.2.18)

| (51)Int.Cl. ⁷ | 識別記号 | F I | テマート*(参考) |
|------------------------------|-------|---------------|-------------------|
| G 0 6 F 17/22 | | G 0 6 F 15/20 | 5 2 4 Z 5 B 0 0 9 |
| 3/00 | 6 5 3 | 3/00 | 6 5 3 A 5 B 0 2 0 |
| 3/02 | 3 6 0 | 3/02 | 3 6 0 G |
| | | 15/20 | 5 0 2 Z |
| | | | 5 2 6 Z |
| 審査請求 未請求 請求項の数8 O L (全 21 頁) | | | |

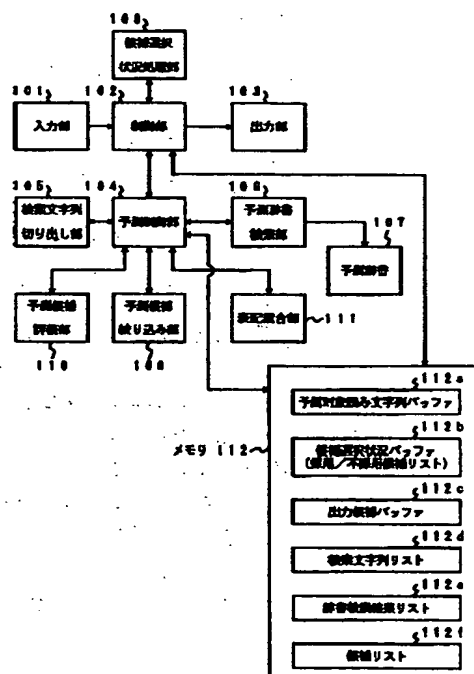
| | | | |
|----------|-----------------------|---------|---|
| (21)出願番号 | 特願平10-211118 | (71)出願人 | 000003078 株式会社東芝 神奈川県川崎市幸区堀川町72番地 |
| (22)出願日 | 平成10年7月27日(1998.7.27) | (72)発明者 | 市村 由美 神奈川県川崎市幸区小向東芝町1番地 株 式会社東芝研究開発センター内 |
| | | (74)代理人 | 100058479 弁理士 鈴江 武彦 (外6名) Fターム(参考) 5B009 ME17 MG06 MH02 5B020 BB10 CC11 DD02 GG21 GG22 |

(54) 【発明の名称】 入力予測装置、入力予測方法及び入力予測プログラムを記録した記録媒体

(57) 【要約】

【課題】文字列の入力中に任意の位置で自動的に入力予測を行い、既に表示されている漢字かな表記との整合性やユーザによる候補採用／不採用の指示を考慮し、不要な候補を排除してユーザに提示する。

【解決手段】入力部１０１から入力された文字列を検索文字列切り出し部１０５に渡して最後尾の文字を含む部分文字列群を生成する。この部分文字列群を予測辞書検索部１０６に渡して予測候補を得、これを出力部１０３により提示する。ここで、提示された予測候補が採用されなかった場合に、予測制御部１０４は当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は当該予測候補の出力を禁止する。また、既に表示されている文字列の表記と矛盾する予測候補の出力を禁止する。これにより、不要な候補を排除してユーザに提示することが可能となる。



【特許請求の範囲】

【請求項1】 文字列を入力する入力手段と、

この入力手段によって入力された文字列から最後尾の文字を含む部分文字列群を生成する部分文字列生成手段と、

各単語の部分文字列とそれに対応する見出しが登録された辞書と、

上記部分文字列生成手段によって生成された部分文字列群を検索文字列として用い、上記辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索する辞書検索手段と、

この辞書検索手段によって得られた予測候補を提示する候補提示手段と、

この候補提示手段によって提示された予測候補が採用されなかった場合に、当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は当該予測候補の出力を禁止する候補制御手段とを具備したことを特徴とする入力予測装置。

【請求項2】 上記候補提示手段によって提示された予測候補を不採用とする指示を行う指示手段を有し、上記候補制御手段は、この指示手段によって不採用の指示がなされた予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は上記不採用の指示がなされた予測候補の出力を禁止することを特徴とする請求項1記載の入力予測装置。

【請求項3】 上記候補提示手段によって提示された予測候補を採用とする指示を行う指示手段を有し、上記候補制御手段は、この指示手段によって採用の指示がなされなかった予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は上記採用の指示がなされなかった予測候補の出力を禁止することを特徴とする請求項1記載の入力予測装置。

【請求項4】 文字列を入力する入力手段と、この入力手段によって入力された文字列から最後尾の文字を含む部分文字列群を生成する部分文字列生成手段と、

各単語の部分文字列とそれに対応する見出しが登録された辞書と、

上記部分文字列生成手段によって生成された部分文字列群を検索文字列として用い、上記辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索する辞書検索手段と、

この辞書検索手段によって得られた予測候補を提示する候補提示手段と、

この候補提示手段によって予測候補を提示する際に、既に表示されている文字列の表記と矛盾する予測候補の出力を禁止する候補制御手段とを具備したことを特徴とする入力予測装置。

【請求項5】 入力された文字列から最後尾の文字を含む部分文字列群を生成し、

この部分文字列群を検索文字列として用い、各単語の部分文字列とそれに対応する見出しが登録された辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索し、

この辞書検索によって得られた予測候補を提示するものであって、

提示された予測候補が採用されなかった場合に、当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は当該予測候補の出力を禁止するようにしたことを特徴とする入力予測方法。

【請求項6】 入力された文字列から最後尾の文字を含む部分文字列群を生成し、

この部分文字列群を検索文字列として用い、各単語の部分文字列とそれに対応する見出しが登録された辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索し、

この辞書検索によって得られた予測候補を提示するものであって、

予測候補を提示する際に、既に表示されている文字列の表記と矛盾する予測候補の出力を禁止するようにしたことを特徴とする入力予測方法。

【請求項7】 入力された文字列を元に、それに続く文字列を予測するための入力予測プログラムを記録した記録媒体であって、

入力文字列から最後尾の文字を含む部分文字列群を生成する手順と、

この部分文字列群を検索文字列として用い、各単語の部分文字列とそれに対応する見出しが登録された辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索する手順と、

この辞書検索によって得られた予測候補を提示する手順と、

提示された予測候補が採用されなかった場合に、当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は当該予測候補の出力を禁止する手順とをコンピュータに実行させるプログラムを記録したコンピュータ読取り可能な記録媒体。

【請求項8】 入力された文字列を元に、それに続く文字列を予測するための入力予測プログラムを記録した記録媒体であって、

入力文字列から最後尾の文字を含む部分文字列群を生成する手順と、

この部分文字列群を検索文字列として用い、各単語の部分文字列とそれに対応する見出しが登録された辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索する手順と、

この辞書検索によって得られた予測候補を提示する手順と、

予測候補を提示する際に、既に表示されている文字列の表記と矛盾する予測候補の出力を禁止する手順とをコン

コンピュータに実行させるプログラムを記録したコンピュータ読取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、パーソナルコンピュータやワードプロセッサ等で日本語文章を作成するために使用される日本語入力システムにおいて、ユーザが入力した文字列を元に、それに続いてユーザが入力しようとする文字列を予測する入力予測装置に係り、特に既に表示されている漢字かな表記との整合性やユーザによる候補採用／不採用の指示を考慮した入力予測を行う入力予測装置と、この装置に用いられる入力予測方法及び入力予測プログラムを記録した記録媒体に関する。

【0002】

【従来の技術】近年、ネットワークが急速に整備され、計算機を用いたコミュニケーションが一般化しつつある。それに伴い、デスクトップ計算機を用いたビジネス文書の作成、論文の執筆といった、文書作成の従来からの主な用途に加えて、携帯情報機器を利用した電子メールの利用、メモの作成といった、より個人的な場面でも計算機で日本語文章を入力する機会が増大している。このような状況の中で、キーボード操作に熟練していない人でも、少ないキータッチで情報を正確に入力するための技術が求められている。

【0003】一方、コンピュータを初めとする種々の電子装置は小型化し、携帯機器の普及も著しい。携帯機器においては、入力操作部のサイズが小さいことから、文書の入力作業の負担はさらに大きく、ユーザの入力に関わる負担を軽減することが望まれる。

【0004】このような要請に対する一つの解決方法として、入力したい熟語の読み文字列の先頭を入力し、それをキーとして、ユーザが望む熟語を検索して提示する文書入力装置（特開昭60-105029号公報）や、変換確定された単語に接続する単語を検索提示するワードプロセッサ（特開平1-260568号公報）、入力したい読み文字列の先頭を入力し、それをキーとして読みを検索して提示するかな漢字変換システム（特開平8-147289号公報）などが提案されている。

【0005】特開昭60-105029号公報では、熟語の読みの頻度情報を記憶し、これによって候補の順位付けを行っている。特開平1-260568号公報では、過去の入力履歴により候補の優先度付けを行っている。特開平8-147289号公報では、変換候補の単語の頻度情報を記憶し、これによって候補の優先度付けを行っている。

【0006】

【発明が解決しようとする課題】上記したように、従来、ユーザの入力負担を軽減するための種々の方法が提案されている。しかしながら、これらの方法はいずれも、ユーザが望む文字列の先頭が入力されることを前提

とし、それをキーにして辞書から前方一致する文字列を予測候補として抽出する仕組みとなっており、文章の入力中に任意の位置で、それに続く文字列を自動的に予測できるような機能は備えていない。

【0007】このため、一般的な連文節入力方式（複数の文節を連続して入力する方式）を用いた場合での入力予測では、入力の途中で一度変換キーを押して入力操作を中断した後、改めて入力したい文字列の先頭を入力するか、あるいは、予測を行いたい位置で予測検索を明示的に指示するなどのキー操作が必要となり、操作が繁雑になる問題がある。

【0008】また、これらの方法はいずれも、入力された文字列が辞書に登録されている文字列の先頭に一致したら、必ずそれを予測候補として提示する。したがって、予測検索を命令せずに自動的に入力予測を行うと、誤った予測を行う可能性が高い。また、学習等により辞書に登録されている表現が増えるに連れ、予測候補が大量に提示され、その中から所望の候補を探して指定する操作が繁雑になる可能性が高い。

【0009】なお、入力中の任意の位置から自動的に入力予測を行うためには、既に表示されている漢字かな表記との整合性や、ユーザによる候補採用／不採用の指示を考慮した表示方法の開発が必要である。

【0010】本発明は上記のような点に鑑みなされたもので、文字列の入力中に任意の位置で自動的に入力予測を行い、既に表示されている漢字かな表記との整合性やユーザによる候補採用／不採用の指示を考慮し、不要な候補を排除してユーザに提示することにより、文書入力の効率を向上させ、入力操作の負担を一層軽減することのできる入力予測装置、入力予測方法及び入力予測プログラムを記録した記録媒体を提供することを目的とする。

【0011】

【課題を解決するための手段】本発明の入力予測装置は、文字列を入力する入力手段と、この入力手段によって入力された文字列から最後尾の文字を含む部分文字列群を生成する部分文字列生成手段と、各単語の部分文字列とそれに対応する見出しが登録された辞書と、上記部分文字列生成手段によって生成された部分文字列群を検索文字列として用い、上記辞書から部分文字列が前方一致あるいは完全一致する見出しを予測候補として検索する辞書検索手段と、この辞書検索手段によって得られた予測候補を提示する候補提示手段と、この候補提示手段によって提示された予測候補が採用されなかった場合に、当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は当該予測候補の出力を禁止する候補制御手段とを具備したことを特徴とする（請求項1）。

【0012】このような構成によれば、辞書検索の結果として提示された予測候補が採用されなかった場合に、

当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は、当該予測候補の出力が禁止される。したがって、採用しなかった候補が次の文字を入力した際に続けて提示されるといった煩わしさが解消される。

【0013】この場合、提示された予測候補を不採用とする指示を行うようにし、この不採用の指示がなされた予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は上記不採用の指示がなされた予測候補の出力を禁止することで（請求項2）、不採用として指示した候補が次の文字を入力した際に続けて提示されることを回避できる。

【0014】また、提示された予測候補を採用とする指示を行うようにし、この採用の指示がなされなかった予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は上記採用の指示がなされなかった予測候補の出力を禁止することで（請求項3）、明示的に不採用の指示を行わなくとも、意図しない候補が次の文字を入力した際に続けて提示されることを回避できる。

【0015】さらに、辞書検索によって得られた予測候補を提示する際に、既に表示されている文字列の表記と矛盾する予測候補の出力を禁止することで（請求項4）、入力補完された文字列を用いて入力予測を行う場合などにおいて、既に表示されている文字列の表記と矛盾しない予測候補のみを提示することができる。

【0016】

【発明の実施の形態】以下、図面を参照して本発明の一実施形態を説明する。図1は本発明の一実施形態に係る入力予測装置の構成を示すブロック図である。なお、本実施形態における入力予測装置は、例えば磁気ディスク等の記録媒体に記録されたプログラムを読み込み、このプログラムによって動作が制御されるコンピュータによって実現される。

【0017】図1において、入力手段としての入力部101は、例えばキーボード、マウス、ペン入力装置、音声入力装置などからなり、文字列の入力や、カーソルの移動、文字列の挿入・削除などの編集指示、次候補表示、候補選択などのコマンド入力を取り込む。

【0018】制御部102は、入力部101が取り込んだ情報を解析して、各処理部へ必要な情報を送る。各処理部での処理結果は再び制御部102に返され、必要な情報は出力部103を介して出力され、例えばディスプレイに表示される。

【0019】ここで、入力部101から入力される各種の情報のうち、読み文字列は制御部102を介して予測制御部104に送られる。予測制御部104では、取り込んだ情報を解析して各処理部へ必要な情報を送る。各処理部での処理結果は再び予測制御部104に返され、予測された読み文字列や漢字かな表記文字列など、必要

な情報は制御部102に返される。さらに、予測候補を提示した場合には、提示された候補の情報と、もし採用された候補があった場合はその候補番号が制御部102を介して候補選択状況処理部108に送られる。

【0020】候補選択状況処理部108は、それらの情報を候補選択状況バッファ112bに記憶し、再び制御部102に返す。受け取った情報は予測制御部104に送られ、候補生成の際に、表記照合部111で参照される。

【0021】メモリ112は、各処理の結果を一時的に保持するための記憶領域であり、例えばRAMなどからなる。このメモリ112には、予測対象読み文字列バッファ112a、候補選択状況バッファ112b、出力候補バッファ112c、検索文字列リスト112d、辞書検索結果リスト112e、候補リスト112fなどが設けられている。

【0022】予測対象読み文字列バッファ112aは、入力された読み文字列を予測対象として格納する。候補選択状況バッファ112bは、提示された予測候補の選択状況を格納するものであって、採用されなかった候補を格納するための不採用候補リストと、採用された候補を格納するための採用候補リストを有する（図15参照）。出力候補バッファ112cは、最終的にユーザに提示すべき予測候補を格納する。

【0023】検索文字列リスト112dは、入力された読み文字列から作成される部分読み文字列を検索用の文字列として格納する（図9参照）。辞書検索結果リスト112eは、予測辞書検索部106によって辞書検索結果として得られた予測候補を格納する。候補リスト112fは、候補選択状況処理部108、予測候補絞り込み部109、予測候補評価部110、表記照合部111によって処理された予測候補を格納する（図10～図13参照）。

【0024】次に、予測制御部104に関わる処理について説明する。予測制御部104は、制御部102を介して、入力あるいは予測補完された読み文字列と候補選択状況を受け取る。この読み文字列は、まず、検索文字列切り出し部105に送られる。検索文字列切り出し部105は、検索用の文字列群を生成し、その結果を予測制御部104に返す。

【0025】この検索用の文字列群は、入力読み文字列の最後尾の文字を含む部分文字列の集合である。例えば、「よろしくお」といった読み文字列の入力では、「よろしくお」に加えて、「ろしくお」、「しくお」、「くお」といったように、先頭から1文字ずつ短くした部分文字列を作成することである。この検索文字列切り出し部105の具体的な処理については後に詳述する。

【0026】予測制御部104は、検索文字列切り出し部105から受け取った検索用の文字列群を予測辞書検索部106に送る。予測辞書検索部106は、予測辞書

107を参照しながら、検索用の読み文字列で辞書の読みと前方一致あるいは完全一致する見出しの読み、漢字かな表記、その他付属する各種情報、例えば品詞や頻度情報などを取り出して、予測制御部104に返す。

【0027】次に、予測制御部104は、予測辞書検索部106から得られた見出しを予測候補評価部110に送る。予測候補評価部110は、送られてきた見出しに対し、付属する各種情報により各種評価点を算出・付与し、さらにそれらを代表する代表評価点を算出・付与して予測制御部104に返す。

【0028】次に、予測制御部104は各種評価点を予測候補絞り込み部109に送る。予測候補絞り込み部109は、各種評価点に対して設定可能な閾値を参照し、各種評価点が設定されている閾値以上の候補のみを選別して、判定結果を予測制御部104に返す。

【0029】さらに、予測制御部104は、予測候補絞り込み部109からの判定結果が「○」であった候補を、表記照合部111に送る。表記照合部111は、候補の読みと表記と、候補選択状況バッファ112bに記憶される採用候補リスト、不採用候補リストとを照合して、条件に適合する候補のみを選別して、判定結果を予測制御部104に返す。予測制御部104では、予測候補絞り込み部109と表記照合部111の判定条件を両方満たした候補を代表評価点順にソートし、上位M個(Mは任意の正の整数)の候補を出力候補バッファ112cにセットし、制御部102に送る。

【0030】以上が本システムの概要である。次に、各処理の詳細についてフローチャートを用いて説明する。

(a) 制御部102の処理動作

まず、制御部102の処理動作について説明する。

【0031】図2は同実施形態における制御部102の処理動作を示すフローチャートである。ステップS201において、制御部102は、メモリ112に設けられた予測対象読み文字列バッファ112a、候補選択状況バッファ112bを初期化した後、以下のような処理を実行する。

【0032】すなわち、ステップS202で、ユーザの入力を待ち、入力があったらステップS203に進む。ステップS203で、その入力情報が読み文字かどうかを判定し、読み文字である場合はステップS204に進み、そうでない場合はステップS210に進む。なお、入力情報が読み文字であるかどうかの判定は、予め読み文字に関する情報が登録されたテーブルを参照する等の処理で行うことが可能である。ここでは、句読点は読み文字ではないとする。また、入力が所謂ローマ字入力である場合には、読み1文字にあたる入力があるまで、ここでバッファリングを行うものとする。

【0033】ステップS204で、予測候補が表示されているかどうか判定し、表示されている場合はステップS205に進み、そうでない場合はステップS206に進む。

進む。

【0034】ステップS206に進んだ場合は、出力候補バッファ112cの情報を候補選択状況処理部108に送り、起動モード「不採用」で候補選択状況処理部108をスタートさせ、ステップS207に進む。ステップS207で、候補選択状況処理部108から結果を受け取り、ステップS205に進む。ステップS205では、入力された読み1文字を予測対象読み文字列バッファ112aに追加し、ステップS208に進む。ステップS208で、予測対象読み文字列バッファ112aの情報と候補選択状況バッファ112bの情報を予測制御部104に送り、入力予測をスタートさせ、ステップS209に進む。

【0035】ステップS209で、予測制御部104から出力候補バッファ112cの情報を受け取り、表示処理を行い、ステップS202に戻り、次のユーザ入力を待つ。表示処理については後に詳述する。

【0036】ステップS210に進んだ場合は、入力が予測候補選択を指示するキー、例えば選択キーであるかどうかを判定する。選択キーである場合はステップS211に進み、そうでない場合はステップS215に進む。ステップS211に進んだ場合は、出力候補バッファ112cの情報と採用された候補の番号を候補選択状況処理部108に送り、起動モード「採用」で候補選択状況処理部108をスタートさせ、ステップS212に進む。

【0037】ステップS212で、候補選択状況処理部108から結果を受け取り、ステップS213に進む。ステップS213で、予測対象読み文字列バッファ112aに予測補完された読み文字列を追加し、ステップS214に進む。ステップS214で、選択された候補を文書入力領域に表示させるなどの表示処理を行い、ステップS202に戻り、次のユーザ入力を待つ。

【0038】ステップS215に進んだ場合は、入力情報が句読点または改行文字であるかどうか判定する。句読点または改行文字である場合は、ステップS216に進み、そうでない場合は、ステップS218に進み、別途定義されている処理を行う。ステップS216では、候補選択状況バッファ112bをクリアし、ステップS217に進む。ステップS217で、予測対象読み文字列バッファ112aをクリアし、ステップS202に戻り、次のユーザ入力を待つ。

【0039】以上の繰り返しにより、制御部102は、入力あるいは予測補完された読み文字列の情報と候補選択の状況を予測制御部104へ、出力候補とユーザの採用／不採用指示情報を候補選択状況処理部108へと受け渡し、結果を受け取って次の処理へと進む。

【0040】なお、上記実施形態では、句読点または改行文字の入力があったときに、候補選択状況バッファ112bをクリアしているが、他の特定記号やキーの入力

を条件にして候補選択状況バッファ112bをクリアするようにしてもよい。

【0041】(b)予測制御部104の処理動作次に、予測制御部104の処理動作について説明する。図3は同実施形態における予測制御部104の処理動作を示すフローチャートである。まず、ステップS301において、予測制御部104は、制御部102から予測対象読み文字列バッファ112aの情報と候補選択状況バッファ112bの情報を受け取った後、以下のような処理を実行する。

【0042】すなわち、ステップS302において、受け取った予測対象読み文字列バッファ112aの情報を検索文字列切り出し部105に送る。検索文字列切り出し部105は検索文字列リスト112dを返すので、ステップS303でその情報を受け取り、ステップS304に進む。

【0043】ここで、検索文字列リスト112dについて説明する。図9に検索文字列リスト112dのフォーマットと処理結果の格納例を示す。検索文字列リスト112dは、検索文字列切り出し部105によって作成される部分文字列群を記憶したものであって、「番号」と「部分読み文字列」からなっている。

【0044】例えば、「しんに」という3文字の予測対象読み文字列に対しては、図9(a)に示すように、「しんに」、「んに」、「に」といったように、最後尾の文字「に」を含む3つの部分読み文字列が検索文字列リスト112dに作成される。

【0045】同様に、予測対象読み文字列「しんにゅ」に対しては、図9(b)に示すように、「しんにゅ」、「んにゅ」、「にゅ」といった4つの部分読み文字列が検索文字列リスト112dに作成される。

【0046】また、予測対象読み文字列「しんにゅうしゃいんけ」に対しては、図9(c)に示すように、「しんにゅうしゃいんけ」、「んにゅうしゃいんけ」、「にゅうしゃいんけ」、「ゅうしゃいんけ」、「うしゃいんけ」、「しゃいんけ」、「ゃいんけ」、「いんけ」、「んけ」、「け」という10個の部分読み文字列が検索文字列リスト112dに作成される。

【0047】なお、この検索文字列切り出し部105の具体的な処理動作については後に詳述する。図3の説明に戻り、ステップS304で、検索文字列リスト112dの件数を変数Rにセットし、ステップS305に進む。ステップS305で、検索文字列リスト112dのリスト番号を示す変数iの値を1にセットすると共に、候補リスト112fのリスト番号を示す変数CNTの値を1にセットし、ステップS306に進む。ステップS306で、 $i \leq R$ が成立するかどうかを判定し、 $i \leq R$ が成立する場合はステップS307に進み、成立しない場合はステップS312に進む。

【0048】ステップS307に進んだ場合は、検索文

字列リスト112dのi番目を予測辞書検索部106に送り、ステップS308に進む。予測辞書検索部106は、予測辞書107を参照して、読みと前方一致あるいは完全一致する見出しを検索するので、ステップS308で、予測辞書検索部106から辞書検索結果リスト112eを受け取り、ステップS309に進む。

【0049】ここで、予測辞書107の構成について説明する。図14に予測辞書107に登録されている情報の例を示す。予測辞書107は、日本語文の一部を構成する漢字かな表記およびその読み文字列などの情報を記憶したものであって、ここでは「読み」、「漢字かな表記」、「品詞」、「表記頻度」、「読み頻度」、「読み分割可能位置」、「表記分割可能位置」からなる。

【0050】例えば、「新入社員」という表記(見出し)に対しては、「しんにゅうしゃいん」という読み、「名詞」という品詞が付与されている。また、表記の頻度を表す値として「53」、読みの頻度を表す値として「35058, 2181, 240, 238, 117, 58, 53, 53」が付与されており、読み分割可能位置を表す値として「1, 6」、表記分割可能位置を表す値として「1, 3」が付与されている。

【0051】ここで、「表記頻度」の「53」は、「新入社員」という表記が使用される頻度を表す数値であり、具体的には、あるコーパス中で「新入社員」という表記の出現した回数などで代表することができる。コーパスとは、多数の文書(テキスト)の集合を言う。

【0052】また、「読み頻度」は、部分読み文字列毎に「,」で区切られており、この例では「しんにゅうしゃいん」という9文字の読みに対して、その部分読み文字列8個に対する8個の数値が付与されている。すなわち、1番目の「35058」は「しん」という部分読み文字列の頻度、2番目の「2181」は「しんに」という部分読み文字列の頻度を表し、以降、「240」は「しんにゅ」、「238」は「しんにゅう」、「117」は「しんにゅうし」、「58」は「しんにゅうしゃ」、「53」は「しんにゅうしゃい」、「53」は「しんにゅうしゃいん」という部分読み文字列の頻度を表している。

【0053】具体的には、あるコーパスに対する読みを付与したコーパス中で、「しん」、「しんに」、「しんにゅ」などの文字列の出現した回数などで代表することができる。

【0054】また、「読み分割可能位置」は、読み文字列を分割できる位置を示すものであって、それぞれ「,」で区切られている。この例では、「しんにゅうしゃいん」という9文字の読みに対して、何文字目から分割して読みを開始できるかを表している。すなわち、1番目の「1」は1文字目から読みを開始できることを、2番目の「6」は、6文字目から読みを開始できることを表している。すなわち、「しんにゅうしゃいん」

といった読み文字列に対しては、1文字目から5文字目までの部分文字列「しんにゅう」と、6文字目から9文字目までの部分文字列「しゃいん」の2つに分割可能であることが示されている。

【0055】また、「表記分割可能位置」は、漢字かな表記（見出し）の文字列を分割できる位置を示すものであって、それぞれ「，」で区切られている。この例では、「新入社員」という4文字の表記に対して、何文字目から分割して表記を開始できるかを示している。すなわち、1番目の「1」は1文字目から表記を開始できることを、2番目の「3」は、3文字目から表記を開始できることを示している。つまり、「新入社員」といった表記に対しては、1文字目から2文字目までの部分文字列「新入」と、3文字目から4文字目までの部分文字列「社員」の2つに分割可能であることが示されている。

【0056】再び、図3の説明に戻り、ステップS309で、辞書検索結果リスト112eの件数を変数Nにセットし、ステップS310に進む。ステップS310で、辞書検索結果リスト112eを候補リスト112fのCNT番目から順にコピーし、ステップS311に進む。ステップS311で、変数CNTをNインクリメント、変数iを1インクリメントしてステップS306に戻る。

【0057】ステップS306からステップS311の繰り返しにより、検索文字列リスト112dに作成された部分読み文字列が辞書検索部106に送られ、図10乃至図13に示すような候補リスト112fが作成される。

【0058】例えば、図9(a)に示す検索文字列リスト112dの場合には、 $i=1$ のとき、部分読み文字列「しんに」が予測辞書検索部106に送られ、予測辞書107の読みと前方一致あるいは完全一致する見出しが検索される。その結果、候補リスト112fには、図10(a)の1番目から3番目の候補がセットされる。この候補リスト112fの「読み」、「表記」、「品詞」、「表記頻度」、「読み頻度」、「読み分割可能位置」、「表記分割可能位置」の各情報は、辞書情報のコピーである。また、「一致長さ」は、辞書検索で一致した読みの長さである。つまり、部分読み文字列「しんに」の文字数3が格納される。

【0059】 $i=2$ のとき、部分読み文字列「んに」が予測辞書検索部106に送られるが、予測辞書107には「んに」と読みが前方一致あるいは完全一致する見出しはないので、候補リスト112fには何もセットされない。

【0060】 $i=3$ のとき、部分読み文字列「に」が予測辞書検索部106に送られ、予測辞書107の読みと前方一致あるいは完全一致する見出しが検索され、その結果、候補リスト112fには、図10(a)の4番目と5番目の候補がセットされる。この場合、「一致長

さ」の情報として、部分読み文字列「に」の文字数1が格納される。

【0061】再び、図3の説明に戻り、ステップS312に進んだ場合は、カレントの候補を示す変数jに1をセットし、ステップS313に進む。ステップS313で、 $j \leq CNT$ が成立するかどうかを判定し、 $j \leq CNT$ が成立する場合はステップS314に進み、成立しない場合はステップS321に進む。

【0062】ステップS314に進んだ場合は、候補リスト112fのj番目を予測候補評価部110に送り、ステップS315に進む。ステップS315で、予測候補評価部110から評価値を受け取り、候補リスト112fに書き込み、それを予測候補絞り込み部109に送り、ステップS316に進む。

【0063】ステップS316で、予測候補絞り込み部109から判定結果を受け取り、候補リスト112fに書き込み、ステップS317に進む。ステップS317で、候補絞り込み判定が「○」かどうかを判定する。

「○」である場合はステップS318に進み、そうでない場合はステップS320に進む。

【0064】ステップS318では、受け取った結果を表記照合部111に送り、ステップS319に進む。ステップS319で、表記照合部111から照合判定結果を受け取り、候補リスト112fに書き込み、ステップS320に進む。ステップS320で、変数jを1インクリメントし、ステップS313に戻る。

【0065】このステップS313からステップS320の繰り返しにより、候補リスト112fに対して、尤度、有用度、評価値、絞り込み判定、表記照合判定が算出・付与される。

【0066】例えば、図10(a)に示す候補リスト112fに対して、図12(a)のような情報が付与された候補リスト112fが作成される。評価値等の具体的な算出方法については後に詳述する。

【0067】ステップS321に進んだ場合は、候補リスト112fのうち、候補絞り込み判定結果と表記照合判定結果が共に「○」であるものを選別し、評価値順にソートし、上位M個(Mは任意の正の整数)の情報を出力候補バッファ112cにセットし、ステップS322に進む。ステップS322で、出力候補バッファ112cの情報を制御部102に返し、ここでの処理を終了する。

【0068】例えば、図12(a)に示す候補リスト112fに対して、 $M=2$ と設定すると、絞り込み判定と表記照合判定が共に「○」である候補は1番目から3番目の候補である。よって、この1番目から3番目の候補が評価値順にソートされ、上位2個の1番目と2番目の候補が出力候補バッファ112cにセットされる。なお、出力候補バッファ112cの形式は候補リスト112fの形式と同様であるのでここでは省略する。

【0069】このように、検索文字列切り出し部105から入力読み文字列に対応する検索用の文字列（部分読み文字列）を取得し、これを予測辞書検索部106に送り、その検索結果を予測候補評価部110、予測候補絞り込み部109、表記照合部111に送って、候補の絞り込みと評価を行い、最終的に評価値の高い候補を出力候補バッファ112cに格納する処理を行う。以上が、予測制御部104の処理の流れである。

【0070】(c) 検索文字列切り出し部105の処理動作

次に、検索文字列切り出し部105の処理動作について説明する。図4は同実施形態における検索文字列切り出し部105の処理動作を示すフローチャートである。まず、ステップS401において、検索文字列切り出し部105は、検索文字列リスト112dを初期化した後、以下のような処理を実行する。

【0071】すなわち、ステップS402で、予測制御部104から予測対象読み文字列バッファ112aの情報を受け取り、ステップS403に進む。ステップS403で、予測対象読み文字列の文字列長を変数LENにセットし、ステップS404に進む。

【0072】ステップS404で、部分文字列として切り出すべき文字位置を示す変数iに1をセットし、ステップS405に進む。ステップS405で、 $i \leq LEN$ が成立するかどうか判定する。 $i \leq LEN$ が成立する場合はステップS406に進み、成立しない場合はステップS408に進む。ステップS406で、予測対象読み文字列のi文字目からLEN文字目までの文字列を検索文字列リスト112dのi番目にコピーし、ステップS407に進む。ステップS407で、変数iを1インクリメントし、ステップS405に戻る。一方、ステップS408に進んだ場合は、検索文字列リスト112dを予測制御部104に返し、ここでの処理を終了する。

【0073】例えば、予測対象読み文字列が「しんにゅうしゃいんけ」であったとする。「しんにゅうしゃいんけ」の尤度： $=100*$

表記頻度/読み頻度の (BASE-1) 番目の値 … (1)

ステップS506では、尤度に0をセットし、ステップS507に進む。ステップS507で、式(2)により有用度（当該候補が有益に使われる度合い）を計算し、ステップS508に進む。

【0080】

評価値： $=10*尤度+表記頻度+有用度$ … (3)

なお、評価値の計算方法については、上記式(3)の他にも、必要に応じて適宜各項目に対する定数倍の値を変えるなど、種々の計算方法が可能である。

【0082】ステップS509で、上記のようにして算出した評価値を予測制御部104に返し、ここでの処理を終了する。例えば、図10(a)の候補リスト112fの場合、j=1のとき、候補リスト112fの1番目

の「け」は10文字なので、変数LENに10がセットされる。i=1のとき、1文字目から10文字目までの部分読み文字列「しんにゅうしゃいんけ」が検索文字列リスト112dの1番目にコピーされる。

【0074】また、i=2のとき、2文字目から10文字目までの部分読み文字列「んにゅうしゃいんけ」が検索文字列リスト112dの2番目にコピーされる。同様にして、i=10まで進み、10文字目の部分読み文字列「け」が、検索文字列リスト112dの10番目にコピーされ、その結果として、図9(c)に示すような検索文字列リスト112dが作成される。

【0075】このように、入力された読み文字列の最後尾の文字を含む部分読み文字列の集合を作成していく。これらの部分読み文字列は辞書検索用の文字列として用いられる。以上が、検索文字列切り出し部105の処理の流れである。

【0076】(d) 予測候補評価部110の処理動作
次に、予測候補評価部110の処理動作について説明する。図5は同実施形態における予測候補評価部110の処理動作を示すフローチャートである。まず、ステップS501において、予測候補評価部110は予測制御部104から候補リスト112fのj番目の候補を受け取り、以下のような処理を実行する。

【0077】すなわち、ステップS502で、候補見出しの読み文字列の長さを変数LENにセットし、ステップS503に進む。ステップS503で、見出しと読み文字列の一致長さを変数BASEにセットし、ステップS504に進む。

【0078】ステップS504で、 $BASE \geq 2$ が成立するかどうかを判定する。 $BASE \geq 2$ が成立する場合はステップS505に進み、成立しない場合はステップS506に進む。ステップS505では、式(1)により尤度（当該候補の尤もらしさの度合い）を計算し、ステップS507に進む。

【0079】

有用度： $=LEN-BASE$ … (2)

ステップS508で、式(3)により、評価値を計算し、ステップS509に進む。

【0081】

の「侵入」が候補評価部110に送られる。この場合、読み文字列の長さは5であるので、変数LENに5がセットされ、変数BASEに一致長さ3がセットされる。

【0083】上記式(1)により尤度を計算する。

尤度： $=100*105/2181$

$=5$

上記式(2)により有用度を計算する。

【0084】

有用度 = $5 - 3$
 $= 2$

上記式(3)により評価値を計算する。

【0085】

評価値 = $10 * 5 + 105 + 2$
 $= 157$

以上の処理により、候補リスト112fの1番目には、図10(b)に示すように、尤度=5、有用度=2、評価値=157といった情報が書き込まれる。

【0086】このように、辞書検索により得られた候補リスト112fの各候補毎に評価値を求め、その結果を予測制御部104に返す。以上が、予測候補評価部110の処理の流れである。

【0087】(e) 予測候補絞り込み部109の処理動作

次に、予測候補絞り込み部109の処理動作について説明する。図6は同実施形態における予測候補絞り込み部109の処理動作を示すフローチャートである。まず、ステップS601において、予測候補絞り込み部109は、予測制御部104から候補リスト112fのj番目を受け取り、以下のような処理を実行する。

【0088】すなわち、ステップS602で、尤度 ≥ 5 が成立するかどうかを判定する。尤度 ≥ 5 が成立する場合は、ステップS603に進み、成立しない場合はステップS605に進む。

【0089】ステップS603に進んだ場合は、有用度 ≥ 2 が成立するかどうかを判定する。有用度 ≥ 2 が成立する場合はステップS604に進み、成立しない場合はステップS605に進む。

【0090】ステップS604では、尤度 ≥ 5 および有用度 ≥ 2 が成立したことにより、その判定結果を「○」としてステップS606に進む。判定結果「○」は予測候補として有効であることを示す。ステップS605では、尤度 ≥ 5 または有用度 ≥ 2 が不成立であったことにより、その判定結果を「×」としてステップS606に進む。判定結果「×」は予測候補として無効であることを示す。なお、ここで用いた尤度の条件値2や有用度の条件値5という数値は、適宜設定を変更して利用可能な値である。

【0091】ステップS606で、各候補に対する判定結果「○」または「×」を予測制御部104に返し、ここでの処理を終了する。例えば、図10(b)の候補リスト112fの場合、j=1のとき、候補リスト112fの1番目の「侵入」が候補絞り込み部109に送られる。この場合、尤度は5、有用度は2であるので、ステップS602、ステップS603の不等式が成立し、判定結果は「○」となる。これにより、候補リスト112fの1番目には、図11(a)に示すような絞り込み判定情報が書き込まれる。

【0092】このように、辞書検索により得られた候補リスト112fの各候補に対し、尤度および有用度の値に基づいて有効か否かを判定し、その判定結果を予測制御部104に返す。以上が、予測候補絞り込み部109の処理の流れである。

【0093】(f) 表記照合部111の処理動作

次に、表記照合部111の処理動作について説明する。図7は同実施形態における表記照合部111の処理動作を示すフローチャートである。まず、ステップS701において、表記照合部111は、予測制御部104から候補リスト112fのj番目を受け取り、以下のような処理を実行する。

【0094】すなわち、ステップS702で、変数CNTに読み分割可能位置数、変数YLENに候補見出しの読み文字列の長さ、変数HLENに候補見出しの表記文字列の長さ、変数「判定」に初期値として「○」をセットし、ステップS703に進む。

【0095】ステップS703で、制御部102から候補選択状況バッファ112b(採用候補リスト、不採用候補リスト)を受け取り、ステップS704に進む。ステップS704で、変数SAIYOに採用候補リストの件数、変数FUSAIYOに不採用候補リストの件数をセットすると共に、変数Sおよび変数Tにそれぞれ1をセットしてステップS705に進む。ステップS705で、 $S \leq FUSAIYO$ が成立するかどうか判定する。 $S \leq FUSAIYO$ が成立する場合は、ステップS706に進み、成立しない場合はステップS708に進む。

【0096】ステップS706に進んだ場合は、不採用候補リストのS番目は候補に含まれるか否かを判定する。候補に含まれる場合は、ステップS719に進み、含まれない場合はステップS707に進む。ステップS707に進んだ場合は、変数Sを1インクリメントし、ステップS705に戻る。一方、ステップS705からステップS708に進んだ場合は、 $T \leq SAIYO$ が成立するかどうかを判定する。 $T \leq SAIYO$ が成立する場合は、ステップS709に進み、成立しない場合は、ステップS719に進む。

【0097】ステップS709に進んだ場合は、変数Mに1、変数YSに1、変数HSに1をセットし、ステップS710に進む。ステップS710で、 $M \leq CNT$ が成立するかどうかを判定する。 $M \leq CNT$ が成立する場合は、ステップS711に進み、成立しない場合はステップS719に進む。

【0098】ステップS711に進んだ場合は、 $M = CNT$ が成立するかどうかを判定する。 $M = CNT$ が成立する場合はステップS712に進み、成立しない場合はS713に進む。ステップS712では、変数YEにYLEN、変数HEにHLENをセットし、ステップS714に進む。ステップS713では、変数YEに「M番目の読み分割可能位置-1」、変数HEに「M番目の表

記分割可能位置-1」をセットし、ステップS714に進む。ステップS714で、YS文字目からYE文字目までの候補見出しの読み文字列が採用候補リストのT番目の読みに含まれるかどうかを判定する。含まれる場合はステップS715に進み、含まれない場合はステップS716に進む。

【0099】ステップS715に進んだ場合は、HS文字目からHE文字目までの候補見出しの表記文字列は採用候補リストのT番目の表記に含まれるかどうかを判定する。含まれる場合はステップS716に進み、含まれない場合はステップS718に進む。ステップS716では、変数YSに「YE+1」、変数HSに「HE+1」をセットし、ステップS717に進む。ステップS717で、変数Tを1インクリメントし、ステップS708に戻る。

【0100】ステップS718に進んだ場合は、判定結果を「×」とし、ステップS719に進む。ステップS719で、判定結果を予測制御部104に返し、ここでの処理を終了する。

【0101】例えば、図11(a)の候補リスト112fの場合、j=1のとき、候補リスト112fの1番目の「侵入」が表記照合部111に送られる。この時点では、候補選択状況バッファ112bは空であるので、ステップS704で、変数SAIYOにも、変数FUSAIYOにも0がセットされる。したがって、ステップS702でセットされる初期値「○」が判定結果として、予測制御部104に返され、候補リスト112fの1番目には、図11(b)に示すような表記照合判定情報が書き込まれる。

【0102】このように、予測候補絞り込み部109からの判定結果が「○」であった候補に対し、その候補の読みと表記と候補選択状況バッファ112bに記憶される採用候補リスト、不採用候補リストとを照合して、条件に適合する候補のみを選別し、その判定結果を予測制御部104に返す処理を行う。以上が、表記照合部111の処理の流れである。

【0103】(g) 候補選択状況処理部108の処理動作

次に、候補選択状況処理部108の処理動作について説明する。図8は同実施形態における候補選択状況処理部108に処理動作を示すフローチャートである。まず、ステップS801において、候補選択状況処理部108は、制御部102から候補選択状況バッファ112b(採用候補リスト、不採用候補リスト)の情報と、出力候補バッファ112cの情報を受け取り、以下のような処理を実行する。

【0104】すなわち、ステップS802で、変数CNTに出力候補バッファ112cの件数、変数SAIYOに採用候補リストの件数、変数FUSAIYOに不採用候補リストの件数をセットし、ステップS803に進

む。

【0105】ステップS803で、候補選択状況処理部108の起動モードが「採用」かどうかを判定する。起動モードが「採用」である場合は、ステップS804に進み、そうでない場合にはステップS808に進む。

【0106】ステップS804に進んだ場合は、制御部102から選択候補番号を受け取り、ステップS805に進む。ステップS805で、変数SELECTに候補番号をセットし、ステップS806に進む。ステップS806で、出力候補バッファ112cのSELECT番目を採用候補リストの(SAIYO+1)番目にコピーし、ステップS807に進む。ステップS807で、候補選択状況バッファ112bの情報を制御部102に送り、ここでの処理を終了する。

【0107】一方、ステップS803からステップS808に進んだ場合には、変数iに1をセットし、ステップS809に進む。そして、ステップS809で、 $i \leq \text{CNT}$ が成立するかどうかを判定し、 $i \leq \text{CNT}$ が成立する場合にはステップS810に進む。

【0108】ステップS810で、出力候補バッファ112cのi番目を不採用候補リストの(FUSAIYO+1)番目にコピーし、ステップS811に進む。ステップS811で、変数iを1インクリメント、変数FUSAIYOを1インクリメントし、ステップS809に戻る。一方、ステップS809で、条件が成立しなかったときは、ステップS807に進み、候補選択状況バッファ112bの情報を制御部102に送り、ここでの処理を終了する。

【0109】このように、候補採用時に、そのとき選択された予測候補の番号を候補選択状況バッファ112bの採用候補リストに登録し、不採用時には、各予測候補の番号を候補選択状況バッファ112bの不採用候補リストに登録する処理を行う。以上が、候補選択状況処理部108の処理の流れである。

【0110】ここで、図15と図16を用いて、予測候補の表示例と候補選択状況バッファ112bのフォーマットと処理結果の格納例、およびそれに関わる制御部102の動きを説明する。

【0111】図16(a)は、制御部102が予測制御部104から予測結果を受け取り、入力された読み情報と予測された漢字かな表記をディスプレイ画面上に表示してユーザに提示した例である。ここでは、「し」、「ん」、「に」と順番に読み文字列が入力され、「しんに」まで入力が進んだ時点で、先に述べたような予測制御部104の処理により、該当する予測候補が出力された場合の例を示している。

【0112】この「しんに」という読み文字列が予測制御部104に送られる。そして、図12(a)に示す候補リスト112fが作成され、予測候補が制御部102に返され、図16(a)のように候補表示が行われる。

なお、図中の下線は入力した読みに対する表記が未確定であることを示している。また、文字列の直後の四角形はカーソルの位置を示している。

【0113】予測候補ウィンドウ1601が開き、そこに2つの予測候補「侵入」、「進入」が表示され、その中の1つ目の候補「侵入」が反転表示されている。ここで、ユーザが次候補キー、例えば下矢印キーを押せば、2つ目の候補「進入」が反転表示される。また、ユーザが候補採用キー、例えばエンターキーを押せば、反転表示されている候補が採用される。

【0114】この例では、表示されている2つの予測候補はどちらも望む候補ではないため、そのどちらの候補も採用せず、続けて、図16(b)のように「ゆ」を入力する。このとき、前の候補表示は取り消される。

【0115】ここで、前回採用しなかった2つの予測候補「侵入」と「進入」が制御部102から候補選択状況処理部108に送られ、これらの候補に関する情報が候補選択状況バッファ112bの不採用候補リストにセットされる。その結果、候補選択状況バッファ112bは図15(a)に示すような状態になる。このとき、候補採用リストは空である。

【0116】続けて、「しんにゆ」という読み文字列と図15(a)に示す候補選択状況バッファ112bの情報が予測制御部104に送られる。これにより、図12(b)に示すような候補リスト112fが作成され、その予測候補が制御部102に返され、図16(b)のように候補表示が行われる。

【0117】この場合、「侵入」と「進入」も候補として候補リスト112fに格納されるが、これらは前回採用されなかったため、その表記照合判定結果は「×」となり、今回の入力に対する候補としては表示されない。

【0118】ここで、ユーザが候補採用キーを押すと、反転表示されている「新入社員」が採用され、図16(c)に示すように、ディスプレイの表示画面が変更される。さらに、このとき採用された候補「新入社員」が制御部102を介して候補選択状況処理部108に送られ、候補選択状況バッファ112bの採用候補リストにセットされる。その結果、候補選択状況バッファ112bは図15(b)に示すような状態になる。

【0119】続けて、図16(d)のように「け」を入力したとする。この入力により、「しんにゅうしゃいんけ」という入力読み文字列および予測補完された読み文字列と、図15(b)の情報が予測制御部104に送られる。そして、図13に示すような候補リスト112fが作成され、その予測候補が制御部102に返され、図16(d)のような候補表示が行われる。

【0120】この場合、「社印ケース入れ」、「社員研修」などの候補も候補リスト112fに格納されるが、「新入社員」といった文字列が既に表示されていることから、「社印ケース入れ」の「社印」のように、「新入

社員」の「社員」と表記が矛盾する候補は排除される。

これにより、「社員研修」が表示されることになる。

【0121】このように、辞書検索の結果として提示された予測候補が採用されなかった場合に、当該予測候補の生成時に参照された文字列を含む文字列に基づいて入力予測を行う間は、当該予測候補の出力が禁止されるため、採用しなかった候補が次の文字を入力した際に続けて提示されるといった煩わしさが解消される。

【0122】さらに、辞書検索によって得られた予測候補を提示する際に、既に表示されている文字列の表記と矛盾する予測候補の出力が禁止されるため、入力補完された文字列を用いて入力予測を行う場合などにおいて、既に表示されている文字列の表記と矛盾しない予測候補のみを提示することができる。

【0123】なお、上記実施形態では、かなで表現される読み文字列をキーボード等で入力する場合を例にとりて説明したが、入力する文字列の形態はかなの読みに限るものではなく、例えばタブレット等の入力装置に文字認識処理を備えたものでは、漢字も直接入力することができるので、この場合は、入力される文字列は漢字かな混じり文字列となる。このような漢字かな混じり文字列を部分文字列として辞書を検索するようにしても良い。

【0124】例えば、手書きで「新」を入力するのは画数が多いため、「しん入」といったような漢字かな混じり文字列の入力から「新入社員」を得たい場合である。このような場合には、図14に示した予測辞書107の読み情報欄をかな文字列に限定せず、漢字かな混じり文字列にまで拡張しておけば、実現可能である。

【0125】要するに、本発明は、上記実施形態に限定されず、要旨を変更しない範囲で適宜変更して実施可能である。また、上述した実施形態において記載した手法は、コンピュータに実行させることのできるプログラムとして、例えば磁気ディスク（フロッピーディスク、ハードディスク等）、光ディスク（CD-ROM、DVD等）、半導体メモリなどの記録媒体に書き込んで各種装置に適用したり、通信媒体により伝送して各種装置に適用することも可能である。本装置を実現するコンピュータは、記録媒体に記録されたプログラムを読み込み、このプログラムによって動作が制御されることにより、上述した処理を実行する。

【0126】

【発明の効果】以上説明したように本発明によれば、文字列の入力中に任意の位置で自動的に入力予測を行い、既に表示されている漢字かな表記との整合性やユーザによる候補採用／不採用の指示を考慮し、不要な候補を排除してユーザに提示することができる。これにより、文書入力の効率が向上し、入力操作の負担を一層軽減して簡易に日本語入力を行うことができるようになる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る入力予測装置の概略

構成を示すブロック図。

【図2】同実施形態における制御部の処理動作を示すフローチャート。

【図3】同実施形態における予測制御部の処理動作を示すフローチャート。

【図4】同実施形態における検索文字列切り出し部の処理動作を示すフローチャート。

【図5】同実施形態における予測候補評価部の処理動作を示すフローチャート。

【図6】同実施形態における予測候補絞り込み部の処理動作を示すフローチャート。

【図7】同実施形態における表記照合部の処理動作を示すフローチャート。

【図8】同実施形態における候補選択状況処理部の処理動作を示すフローチャート

【図9】同実施形態における検索文字列リストのフォーマットと処理結果の格納例を示す図。

【図10】同実施形態における候補リストのフォーマットと処理結果の格納例を示す図。

【図11】同実施形態における候補リストのフォーマットと処理結果の格納例を示す図。

【図12】同実施形態における候補リストのフォーマットと処理結果の格納例を示す図。

【図13】同実施形態における候補リストのフォーマットと処理結果の格納例を示す図。

【図14】同実施形態における予測辞書に登録される情報の例を示す図。

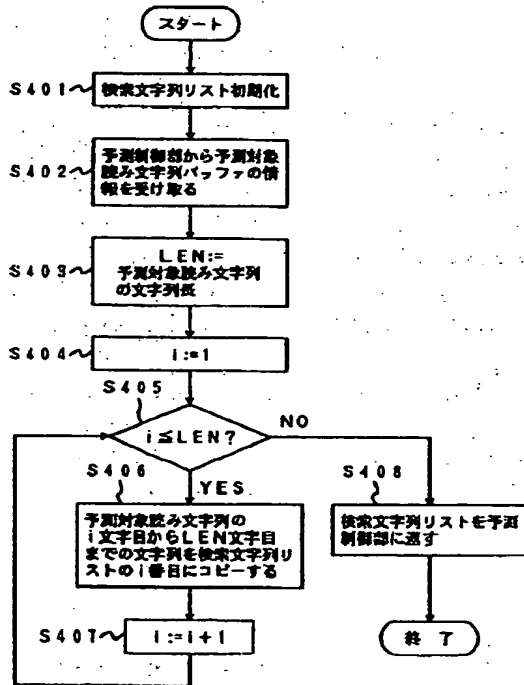
【図15】同実施形態における候補選択状況バッファのフォーマットと処理結果の格納例を示す図。

【図16】同実施形態における文書入力時の画面表示例を示す図。

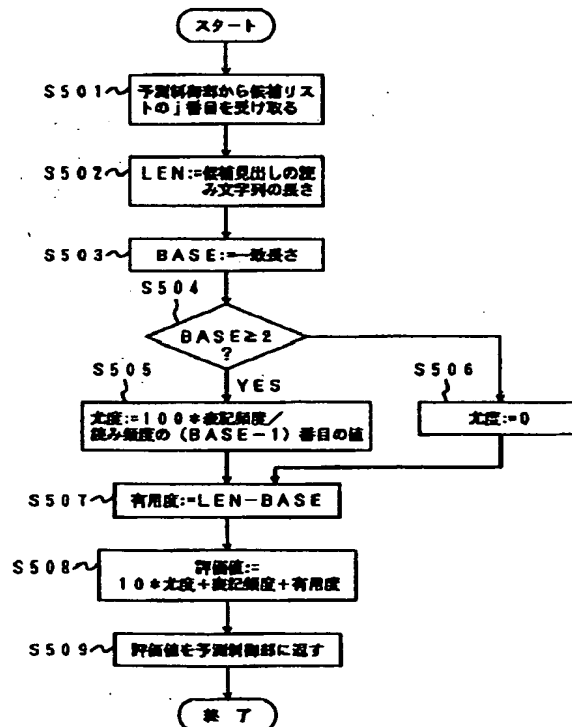
【符号の説明】

101…入力部
102…制御部
103…出力部
104…予測制御部
105…検索文字列切り出し部
106…予測辞書検索部
107…予測辞書
108…候補選択状況処理部
109…予測候補絞り込み部
110…予測候補評価部
111…表記照合部
112…メモリ
112a…予測対象読み文字列バッファ
112b…候補選択状況バッファ
112c…出力候補バッファ
112d…検索文字列リスト
112e…辞書検索結果リスト
112f…候補リスト

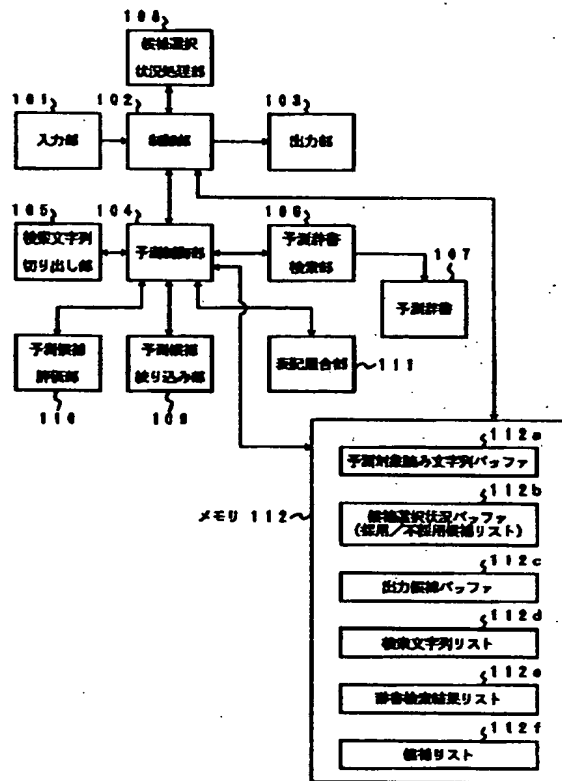
【図4】



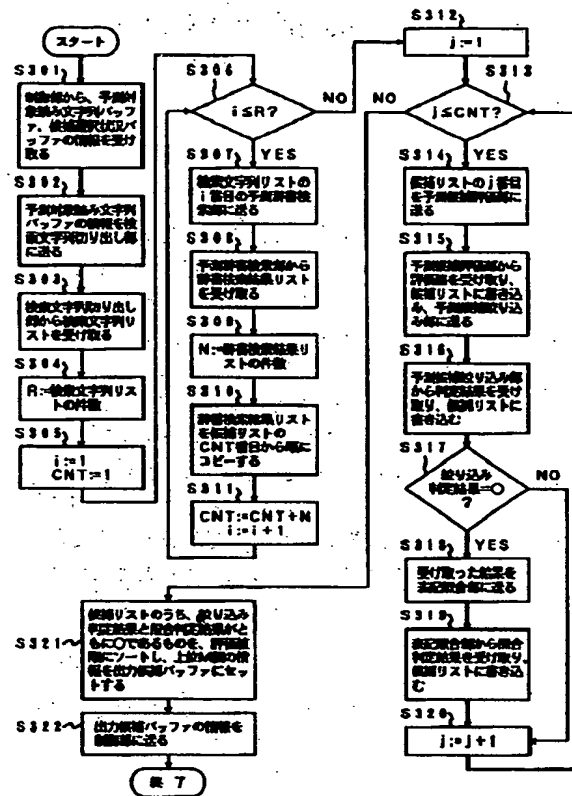
【図5】



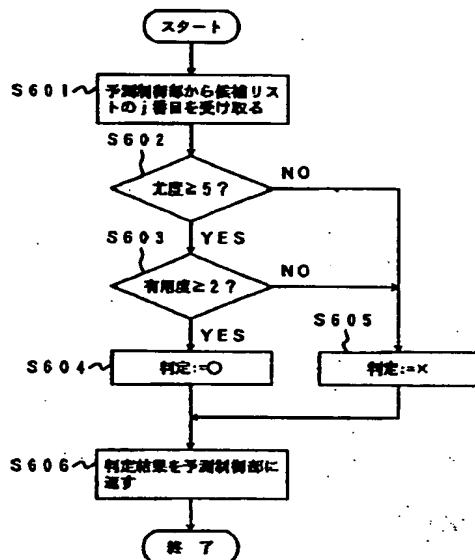
【図 1】



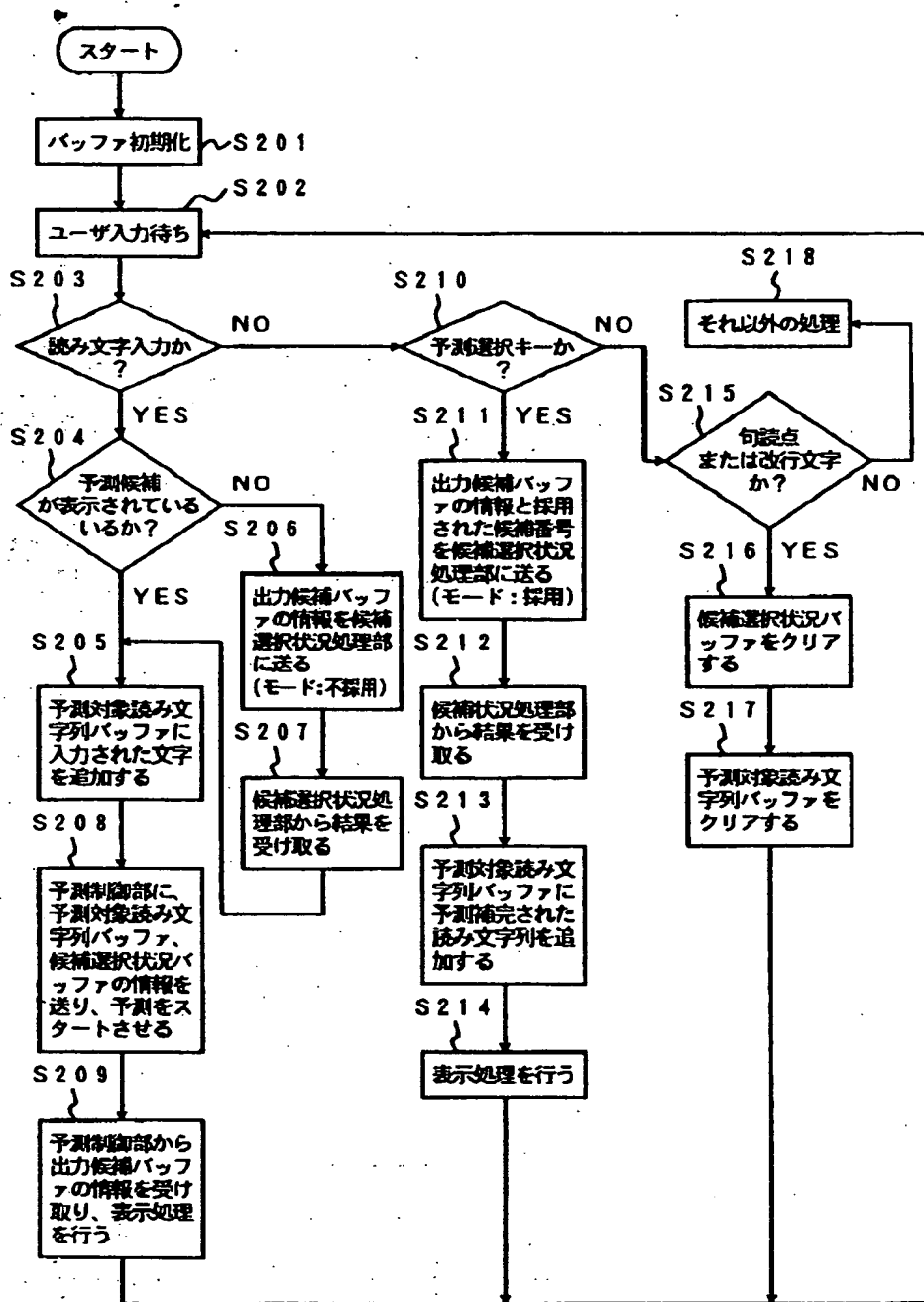
【☒3】



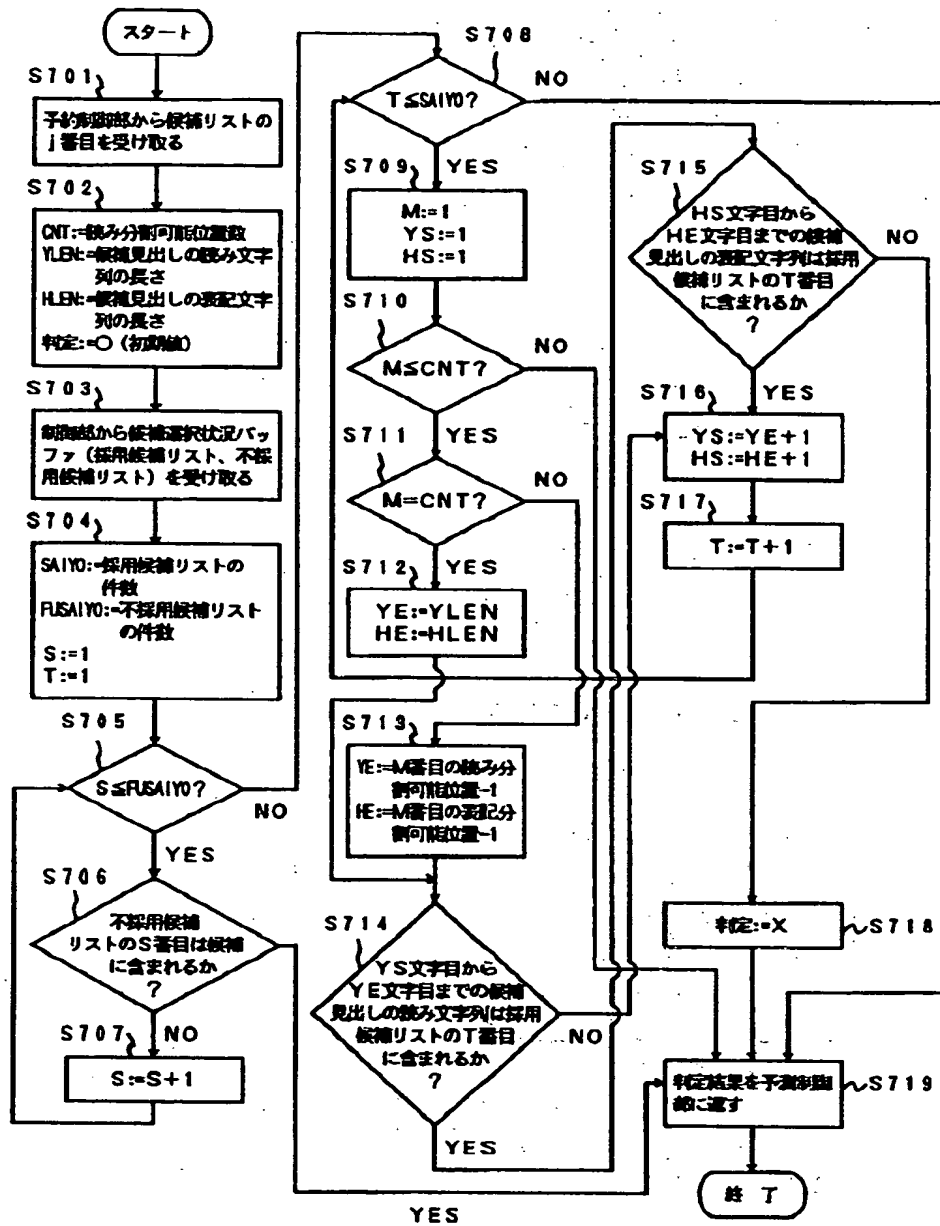
【例 6】



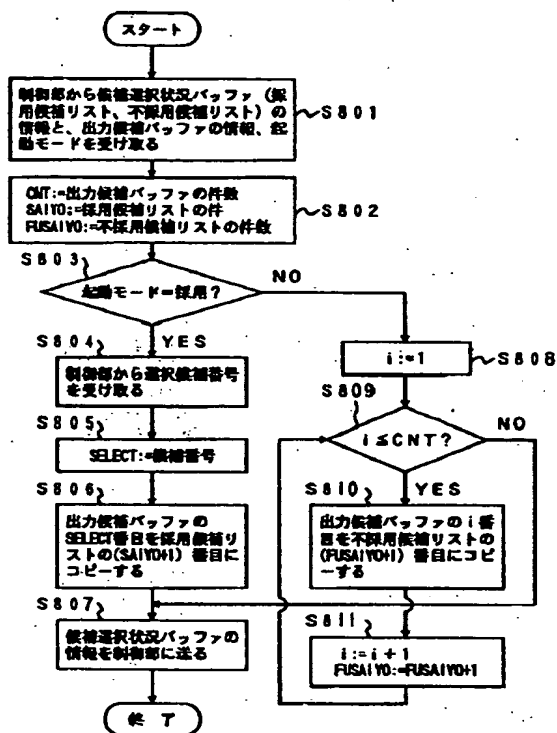
【図2】



【図7】



【図8】



【図9】

検索文字列リスト112d

(a)

| 番号 | 部分読み文字列 |
|----|---------|
| 1 | しんに |
| 2 | んに |
| 3 | に |

(b)

| 番号 | 部分読み文字列 |
|----|---------|
| 1 | しんに |
| 2 | んに |
| 3 | に |
| 4 | |

(c)

| 番号 | 部分読み文字列 |
|----|------------|
| 1 | しんにゅうしゃいんけ |
| 2 | んにゅうしゃいんけ |
| 3 | にゅうしゃいんけ |
| 4 | ゅうしゃいんけ |
| 5 | うしゃいんけ |
| 6 | しゃいんけ |
| 7 | ゃいんけ |
| 8 | いんけ |
| 9 | んけ |
| 10 | け |

【図13】

候補リスト112f

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用度 | 評価値 | 読み 判定 | 表記 適合判定 |
|----|---------------|-------------|------|----------|--|----------------------|----------------------|----------|----|-----|-----|----------|------------|
| 1 | しゃいん げーすいれ | 社印 ケース入れ | 名詞 | 6 | 42707, 1 098, 797, 9, 7, 7, 6, 6 | 1, 5, 8 | 1, 3, 6 | 5 | 67 | 4 | 690 | ○ | × |
| 2 | しゃいん けんしゅう | 社員研修 | サ変名詞 | 6 | 42707, 1 098, 797, 9, 8, 7, 7, 7 | 1, 5 | 1, 3 | 5 | 67 | 4 | 680 | ○ | ○ |
| 3 | けんしき | 見識 | 名詞 | 19 | 32888, 2 480, 22 | 1 | 1 | 1 | 0 | 3 | 22 | × | |
| 4 | けんしゅう | 検収 | サ変名詞 | 12 | 32888, 2 460, 519, 253 | 1 | 1 | 1 | 0 | 4 | 16 | × | |
| 5 | けんしゅう | 研修 | サ変名詞 | 225 | 32888, 2 460, 519, 253 | 1 | 1 | 1 | 0 | 4 | 229 | × | |

【図10】

候補リスト112f

(a)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞り 込み 判定 | 表記 照合 判定 |
|----|---------------|------|------|----------|--|----------------------|----------------------|----------|----|---------|-----|----------------|----------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181.240. 238 | 1 | 1 | 3 | | | | | |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181.240. 238 | 1 | 1 | 3 | | | | | |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181.240. 238, 117. 58, 53, 53 | 1, 6 | 1, 3 | 3 | | | | | |
| 4 | にほん | 日本 | 名詞 | 12925 | 13797.1 2734 | 1 | 1 | 1 | | | | | |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557.411 | 1 | 1 | 1 | | | | | |

(b)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞り 込み 判定 | 表記 照合 判定 |
|----|---------------|------|------|----------|--|----------------------|----------------------|----------|----|---------|-----|----------------|----------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181.240. 238 | 1 | 1 | 3 | 5 | 2 | 157 | | |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181.240. 238 | 1 | 1 | 3 | | | | | |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181.240. 238, 117. 58, 53, 53 | 1, 6 | 1, 3 | 3 | | | | | |
| 4 | にほん | 日本 | 名詞 | 12925 | 13797.1 2734 | 1 | 1 | 1 | | | | | |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557.411 | 1 | 1 | 1 | | | | | |

【図11】

候補リスト112f

(a)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞込み 判定 | 表記照合 判定 |
|----|---------------|------|------|----------|---|----------------------|----------------------|----------|----|---------|-----|-----------|------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181.240. 238 | 1 | 1 | 3 | 5 | 2 | 157 | ○ | |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181.240. 238 | 1 | 1 | 3 | | | | | |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181.240. 238.117. 53.53.53 | 1.6 | 1.3 | 3 | | | | | |
| 4 | にほん | 日本 | 名詞 | 12925 | 13797.1 2734 | 1 | 1 | 1 | | | | | |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557.411 | 1 | 1 | 1 | | | | | |

(b)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞込み 判定 | 表記照合 判定 |
|----|---------------|------|------|----------|---|----------------------|----------------------|----------|----|---------|-----|-----------|------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181.240. 238 | 1 | 1 | 3 | 5 | 2 | 157 | ○ | ○ |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181.240. 238 | 1 | 1 | 3 | | | | | |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181.240. 238.117. 53.53.53 | 1.6 | 1.3 | 3 | | | | | |
| 4 | にほん | 日本 | 名詞 | 12925 | 13797.1 2734 | 1 | 1 | 1 | | | | | |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557.411 | 1 | 1 | 1 | | | | | |

【図12】

候補リスト112f

(a)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞込み 判定 | 表記 照合 判定 |
|----|---------------|------|------|----------|--|----------------------|----------------------|----------|----|---------|-------|-----------|----------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181,240. 238 | 1 | 1 | 3 | 5 | 2 | 157 | ○ | ○ |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181,240. 238 | 1 | 1 | 3 | 5 | 2 | 122 | ○ | ○ |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181,240. 238, 117. 58, 53, 53 | 1,6 | 1,3 | 2 | 5 | 6 | 109 | ○ | ○ |
| 4 | にほん | 日本 | 名詞 | 12325 | 13797.1 2734 | 1 | 1 | 1 | 0 | 2 | 12927 | × | |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557,411 | 1 | 1 | 1 | 0 | 3 | 414 | × | |

(b)

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み 分割 可能 位置 | 表記 分割 可能 位置 | 一致 長さ | 尤度 | 有用 度 | 評価値 | 絞込み 判定 | 表記 照合 判定 |
|----|---------------|------|------|----------|--|----------------------|----------------------|----------|----|---------|-----|-----------|----------------|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058.2 181,240. 238 | 1 | 1 | 4 | 44 | 1 | 546 | ○ | × |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058.2 181,240. 238 | 1 | 1 | 4 | 44 | 1 | 511 | ○ | × |
| 3 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058.2 181,240. 238, 117. 58, 53, 53 | 1,6 | 1,3 | 4 | 22 | 5 | 278 | ○ | ○ |
| 5 | にゅーす | ニュース | 名詞 | 411 | 10554.1 557,411 | 1 | 1 | 2 | 4 | 2 | 453 | × | |

【図14】

予選時書107

| 読み | 表記 | 品名 | 表記順度 | 読み順度 | 読み分割 可能位置 | 表記分割 可能位置 |
|-----------|-------------|------|-------|--|--------------|--------------|
| ... | | | | | | |
| けんしき | 見聞 | 名詞 | 19 | 32888, 2460, 22 | 1 | 1 |
| けんしゅう | 検収 | サ変名詞 | 12 | 32888, 2460, 519, 253 | 1 | 1 |
| けんしゅう | 研修 | サ変名詞 | 225 | 32888, 2460, 519, 253 | 1 | 1 |
| ... | | | | | | |
| しゃいんげーすいれ | 社印ケース 入れ | 名詞 | 6 | 42707, 1098, 797, 9, 7, 7, 6, 6 | 1, 5, 8 | 1, 3, 6 |
| しゃいんけんしゅう | 社員研修 | サ変名詞 | 6 | 42707, 1098, 797, 9, 7, 7, 7, 7 | 1, 5 | 1, 3 |
| しゃかい | 社会 | 名詞 | 2937 | 42707, 4139, 3284 | 1 | 1 |
| しゃがいひ | 社外制 | 名詞 | 2 | 42707, 3293, 207, 2 | 1, 5 | 1, 3 |
| ... | | | | | | |
| しんにゅう | 新入 | サ変名詞 | 105 | 35058, 2181, 240, 238 | 1 | 1 |
| しんにゅう | 進入 | サ変名詞 | 79 | 35058, 2181, 240, 238 | 1 | 1 |
| しんにゅうしゃいん | 新入社員 | 名詞 | 53 | 35058, 2181, 240, 238, 11 7, 58, 53, 53 | 1, 6 | 1, 3 |
| ... | | | | | | |
| にほん | 日本 | 名詞 | 12925 | 13797, 12734 | 1 | 1 |
| にゅーず | ニュース | 名詞 | 411 | 10554, 1557, 411 | 1 | 1 |
| ... | | | | | | |

【図16】

(a) しんにゅう 1601

予選候補ウィンドウ図

1. 新入社員

2. 進入

1/2

(b) しんにゅう 1601

予選候補ウィンドウ図

1. 新入社員

1/1

(c) 新入社員

(d) 新入社員 1601

予選候補ウィンドウ図

1. 新入社員

1/1

【図1.5】

候補選取状況バッファ109b

不採用候補リスト

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み分割 可能位置 | 表記分割 可能位置 | 一致 長さ | 尤度 | 有用度 | 評価値 |
|----|-------|----|------|----------|-------------------------------|--------------|--------------|----------|----|-----|-----|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058, 21 81, 240, 23 8 | 1 | 1 | 3 | 5 | 2 | 157 |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058, 21 81, 240, 23 8 | 1 | 1 | 3 | 5 | 2 | 122 |

採用候補リスト

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み分割 可能位置 | 表記分割 可能位置 | 一致 長さ | 尤度 | 有用度 | 評価値 |
|----|----|----|----|----------|------|--------------|--------------|----------|----|-----|-----|
| | | | | | | | | | | | |

不採用候補リスト

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み分割 可能位置 | 表記分割 可能位置 | 一致 長さ | 尤度 | 有用度 | 評価値 |
|----|-------|----|------|----------|-------------------------------|--------------|--------------|----------|----|-----|-----|
| 1 | しんにゅう | 侵入 | サ変名詞 | 105 | 35058, 21 81, 240, 23 8 | 1 | 1 | 3 | 5 | 2 | 157 |
| 2 | しんにゅう | 進入 | サ変名詞 | 70 | 35058, 21 81, 240, 23 8 | 1 | 1 | 3 | 5 | 2 | 122 |

採用候補リスト

| 番号 | 読み | 表記 | 品名 | 表記 頻度 | 読み頻度 | 読み分割 可能位置 | 表記分割 可能位置 | 一致 長さ | 尤度 | 有用度 | 評価値 |
|----|---------------|------|----|----------|---|--------------|--------------|----------|----|-----|-----|
| 1 | しんにゅう しゃいん | 新入社員 | 名詞 | 53 | 35058, 21 81, 240, 23 8, 117, 58, 53, 63 | 1, 6 | 1, 3 | 4 | 22 | 5 | 278 |